EE 4302 Advanced Control Systems
Part II: Nonlinear System

W.K. Ho
E4-08-13
65166286
elehowk@nus.edu.sg

TOPICS:

Root-Locus Analysis of Nonlinear System

Phase-Plane Analysis

Describing Function Analysis

Linearization

Input-Output Linearization

RECOMMENDED TEXT:

Franklin, Powell and Emami-Naeini, "Feedback Control of Dynamical Systems," 5th Edition, Prentice-Hall 2005.

Slotine and Li, "Applied Nonlinear Control," Prentice-Hall, 1991

Hassan K. Khalil, "Nonlinear System," 3rd Edition, Prentice-Hall, 2002

ROOT LOCUS ANALYSIS OF NONLINEAR SYSTEM

For nonlinearity that has no dynamics and is well approximated as a gain that varies as the size of its input varies, the root locus analysis can be used. Examples are given in the Figure below

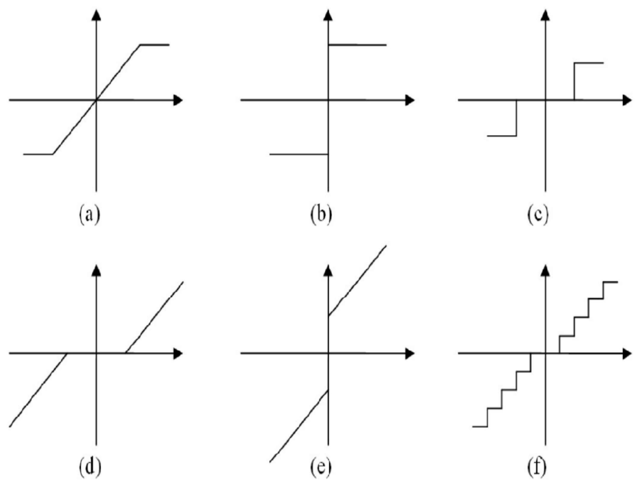


Figure 1: Nonlinear elements with no dynamics: (a) saturation, (b) relay, (c) relay with dead zones, (d) gain withdead zone, (e) preloaded spring and (f) quantization.

Example

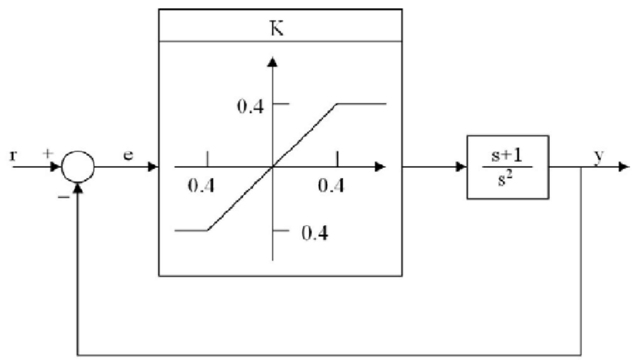


Figure 2: Dynamic system with saturation

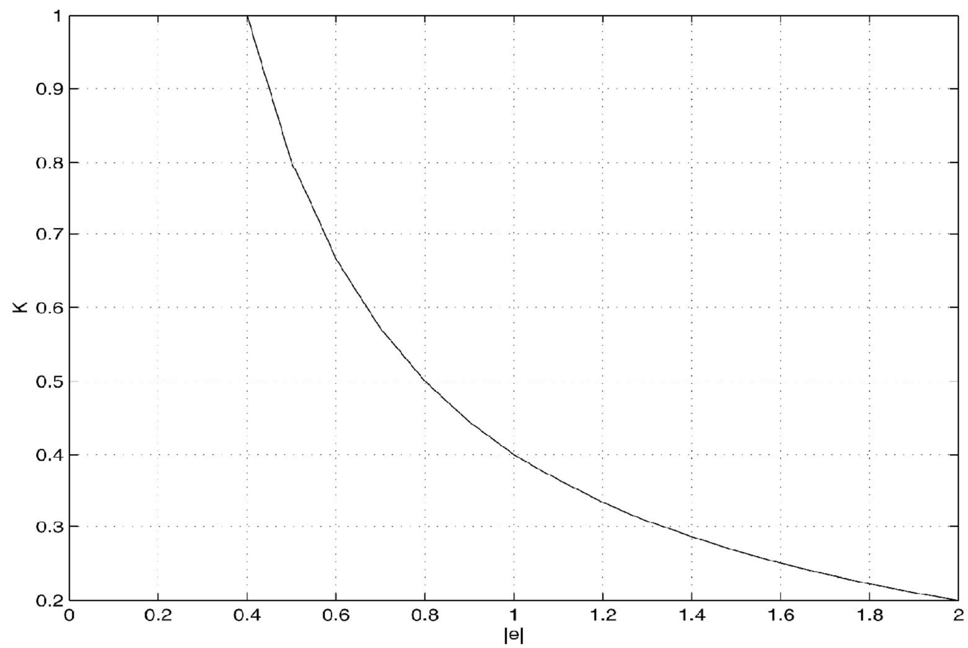


Figure 3: Effective gain of saturation

Table of Roots

K	$s = \frac{-K \pm \sqrt{K^2 - 4K}}{2}$
0	0, 0
0.5	$-0.25 \pm 0.66j$
1	$-0.5 \pm 0.86j$
1.5	$-0.75 \pm 0.97j$
2	$-1 \pm j$
$\vdots$	$\vdots$
∞	$-1, -\infty$

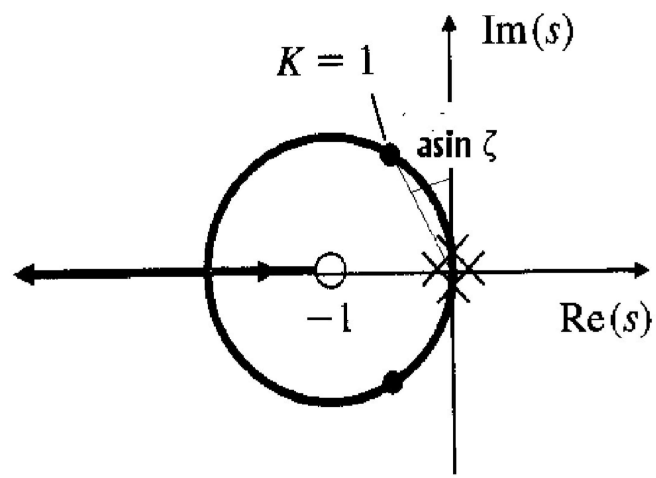


Figure 4: Root locus of $(s+1)/s^3$ with effective gain K.

As the input gets larger, the response has more and more overshoot and slower and slower recovery. This can be explained by noting that larger and larger input signals correspond to smaller and smaller effective gain, K. From the Table of roots and root-locus plot, we see that as K decreases, the closed-loop poles move closer to the origin and have a smaller damping ratio ζ. This results in the longer

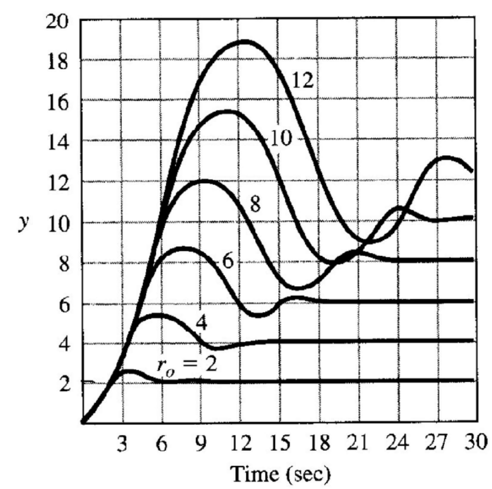


Figure 5: Step responses of system for various input step sizes.

settling times ($t_s = \frac{4}{\zeta\omega_n}$), increased overshoot ($M_p = \exp\left(\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}\right)$) and greater oscillatory response.

<u>Example</u>

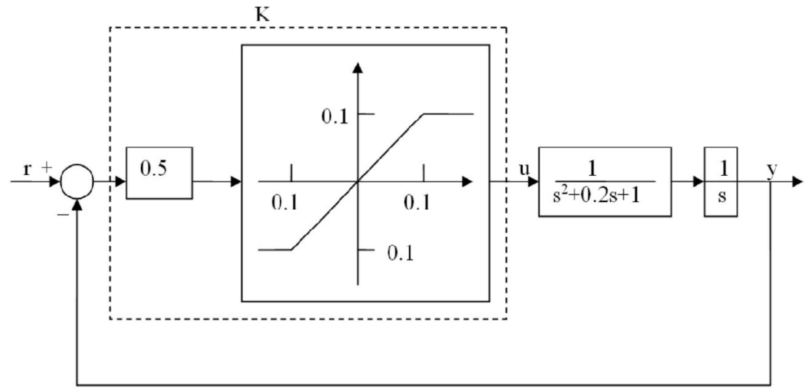


Figure 6: .

K	Roots (Closed-Loop Poles)
0	$0,\ -0.1 \pm 0.99j$
0.2	$-0.2,\ 0 \pm j$
0.5	$-0.45,\ 0.12 \pm 1.05j$
$\vdots$	$\vdots$

From the root locus, the imaginary-axis crossing is at $\omega = 1$, $K =$

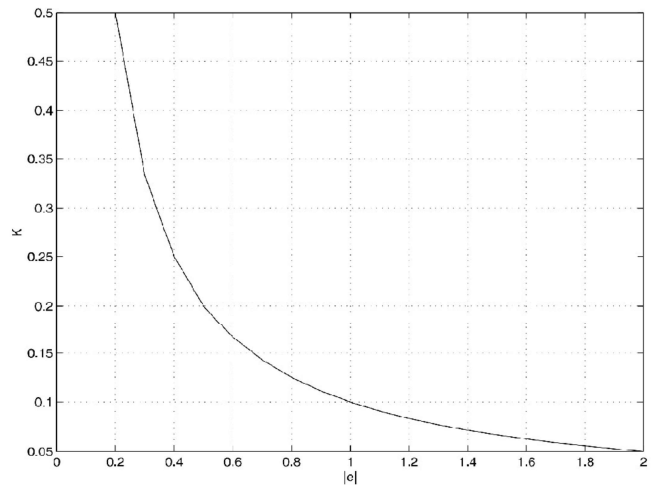


Figure 7: Effective gain of saturation

0.2. With $K = 0.5$ the system is initially unstable but become stable as the gain decreases. The response of the system build up due to the instability until the magnitude is sufficiently large that the effective gain is lowered to $K = 0.2$ and then stop growing but oscillates. The oscillation frequency is $\omega = 1$ rad/sec with constant amplitude. The response is approaching a periodic solution of fixed amplitude known

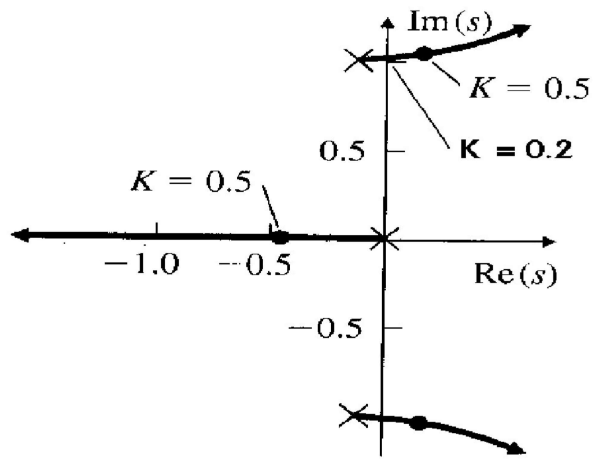


Figure 8: Root locus for the system with effective gain K.

as a limit cycle. Describing function can be used to predict stability and limit-cycle amplitude.

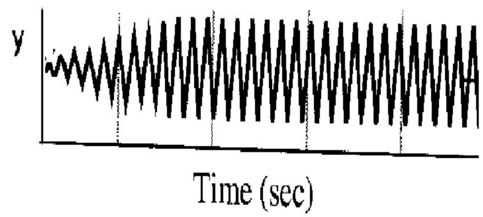


Figure 9: Response of system.

PHASE PLANE ANALYSIS

A graphical method for studying second order systems described by

$$\begin{aligned}\dot{x}_1 &= f_1(x_1, x_2)\\ \dot{x}_2 &= f_2(x_1, x_2)\end{aligned}$$

The advantage is that it allows visual examination. The disadvantage is that it is mainly limited to second-order systems although extensions to third-order systems can often be achieved with the aid of computer graphics.

Example

Consider a mass-spring system:

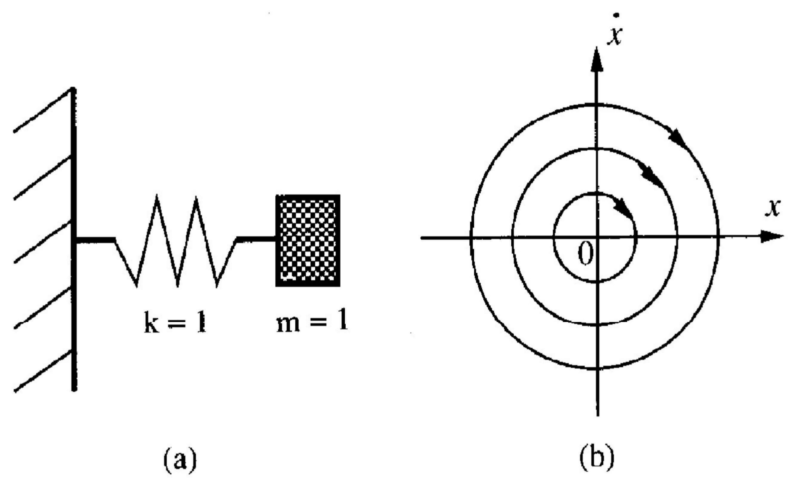


Figure 1: From the phase portrait, we know the position and speed at all time.

$$\ddot{x} + x = 0$$

Set $x_1 = x$, $x_2 = \dot{x}$ gives

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 \end{aligned}$$

Taking Laplace and solving gives

$$
\begin{aligned}
s^2X(s) - sx_0 - \dot{x}_0 + X(s) &= 0 \\
X(s) &= \frac{sx_0}{s^2+1} \\
x(t) &= x_0 \cos t \\
\dot{x}(t) &= -x_0 \sin t \\
x^2(t) + \dot{x}^2(t) &= x_0^2 \\
x_1^2 + x_2^2 &= x_0^2
\end{aligned}
$$

Equilibrium Points

Defined as a point where the system states can stay forever, its location can be obtained from solving $\dot{x}_1 = 0$, $\dot{x}_2 = 0$. i.e. $f_1(x_1, x_2) = 0$, $f_2(x_1, x_2) = 0$ or $\dot{x} = 0$, $\ddot{x} = 0$.

Example

Consider the system

$$\ddot{x} + 0.6\dot{x} + 3x + x^2 = 0$$

Set $\ddot{x} = 0$, $\dot{x} = 0$, give x = 0, -3. The system has two equilibrium points one at (0,0) and the other at (-3,0). The motion patterns of the system trajectories in the vicinity of the two equilibrium points have different natures. The trajectories move towards the points $x = 0$ while moving away from the point $x =$ -3.

Constructing the Phase Portrait

Today, phase portraits are drawn by computers. It is still useful to

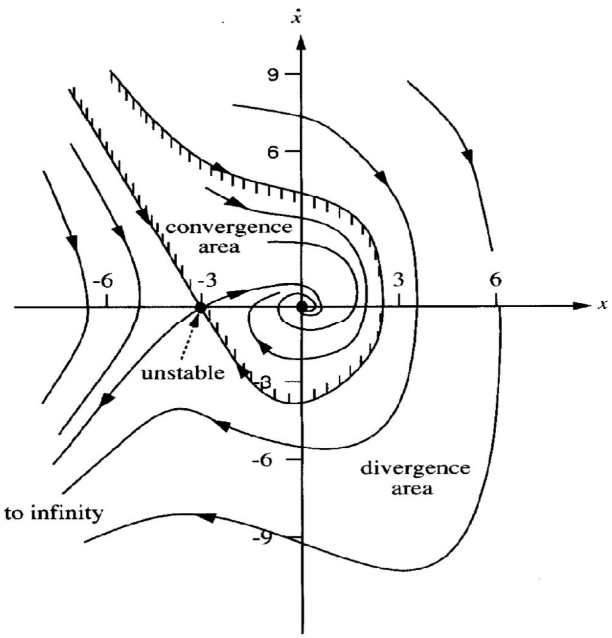


Figure 2: .

learn how to roughly sketch phase portrait or quickly verify computer outputs.

Analytical Method

First technique: solve the differential equation for x_1 and x_2 as a function of time t. We did that for the mass-spring example.

Second technique: use

$$\frac{dx_2}{dx_1} = \frac{f_2(x_1, x_2)}{f_1(x_1, x_2)}$$

Example

Consider the mass-spring system:

$$\ddot{x} + x = 0$$

Set $x_1 = x$, $x_2 = \dot{x}$ gives

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 \end{aligned}$$

$$\frac{dx_2}{dx_1} = -\frac{x_1}{x_2}$$

Integrating gives

$$x_1^2 + x_2^2 = x_0^2$$

Computer Method

The Method of Isoclines

Consider

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_1, x_2) \end{aligned}$$

At a point (x_1, x_2) in the phase plane, the slope (α) of the tangent to the trajectory can be determined by

$$\frac{dx_2}{dx_1} = \frac{f_2}{f_1} = \alpha$$

An isocline is defined to be the locus of the points with a given tangent slope. That is to say that points on the curve

$$f_2 = \alpha f_1$$

all have the same tangent slope α and is an isocline.

In the method of isoclines, the phase portrait of a system is generated in 4 steps: 1) draw the isocline, 2) draw the tangents 3) draw the phase portrait 4) draw the arrows by substituting some values on the phase plane trajectory into the $\dot{x}_1 = f_1$ and $\dot{x}_2 = f_2$ equations to determine whether x_1 and x_2 are increasing or decreasing.

<u>Example</u>

Consider the Van der Pol equation

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = 0$$

Set $x_1 = x$ and $x_2 = \dot{x}$ gives

$$\begin{aligned} \dot{x}_1 &= x_2 = f_1 \\ \dot{x}_2 &= -\mu(x_1^2 - 1)x_2 - x_1 = f_2 \end{aligned}$$

An isocline of slope α is defined by

$$\frac{dx_2}{dx_1} = \frac{f_2}{f_1} = \frac{-\mu(x_1^2 - 1)x_2 - x_1}{x_2} = \alpha$$

Therefore, the points on the curve

$$x_2 = \frac{x_1}{(\mu - \mu x_1^2) - \alpha}$$

all have the same slope α. By taking α of different values, different isoclines can be obtained, as plotted below. Short lines segments are drawn on the isoclines to generate a field of tangent directions. The phase portraits can then be obtained, as shown in the plot. It is interesting to note that there exists a closed curve in the portrait, and the trajectories starting from both outside and inside converge to this curve. This closed curve corresponds to a limit cycle, as reflected by the motion pattern in the Figure below.

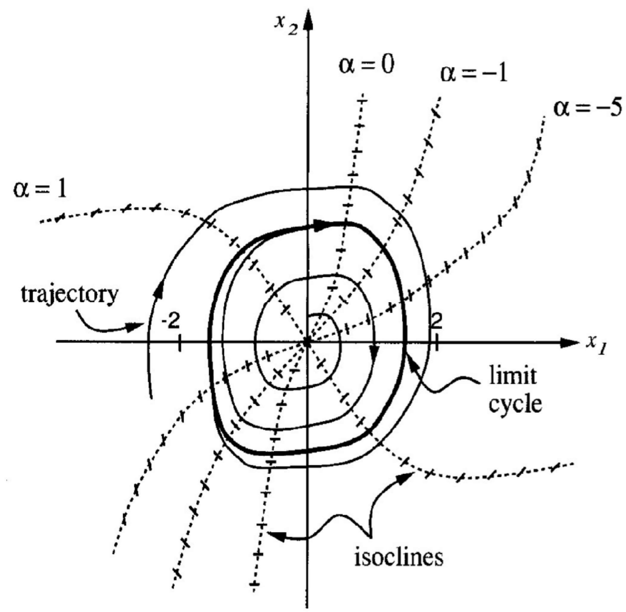


Figure 3: The Method of Isoclines.

Determining Time from Phase Portraits

Method 1

From

$$\frac{\Delta x}{\Delta t} \approx \dot{x}$$
$$\Delta t \approx \frac{\Delta x}{\dot{x}}$$

Divide the trajectory into a number of small segments of Δx, find the time for each segment, and then add up the results.

Method 2

From

$$\dot{x} = \frac{dx}{dt}$$
$$t - t_0 = \int_{x_0}^{x} (1/\dot{x}) dx$$

Plot a phase plane portrait with new coordinates x and $1/\dot{x}$, then the area under the curve is the time interval.

Limit Cycles

The local behaviour of a nonlinear system can be approximated by a linear system behaviour. Yet, nonlinear systems can display much more complicated patterns in the phase plane, such as multiple equilibrium points and limit cycles.

In the phase plane, a limit cycle is defined as an isolated closed curve. Trajectories inside the curve and those outside the curve all tend to this curve, while motion started on this curve will stay on it forever, circling periodically around. The Van der Pol equation has a limit cycle. The mass-spring-damper system do not have limit cycles because surrounding trajectories do not converge to any cycles.

Definitions

- Stable limit cycle: all trajectories in the vicinity of the limit cycle converge to it as $t \rightarrow \infty$.
- Unstable limit cycle: all trajectories in the vicinity of the limit cycle diverge from it as $t \rightarrow \infty$.
- Semi-stable limit cycle: some trajectories converge and some diverge as $t \rightarrow \infty$.

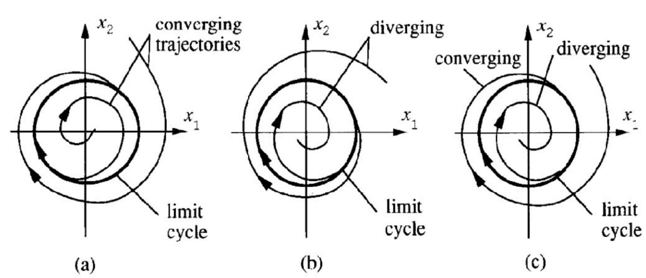


Figure 4: .

DESCRIBING FUNCTION ANALYSIS

For some nonlinear systems and under certain conditions, an extended version of the frequency response method, called the describing function method can be used to approximately analyze and predict nonlinear behaviour. The main use of describing function method is for the prediction of limit cycles in nonlinear systems.

Applications Domain

Simply speaking, any systems which can be transformed into the configuration in the Figure below can be studied using describing functions.

There is an important classes of systems in this category. This important class consists of "almost" linear systems. By "almost" linear

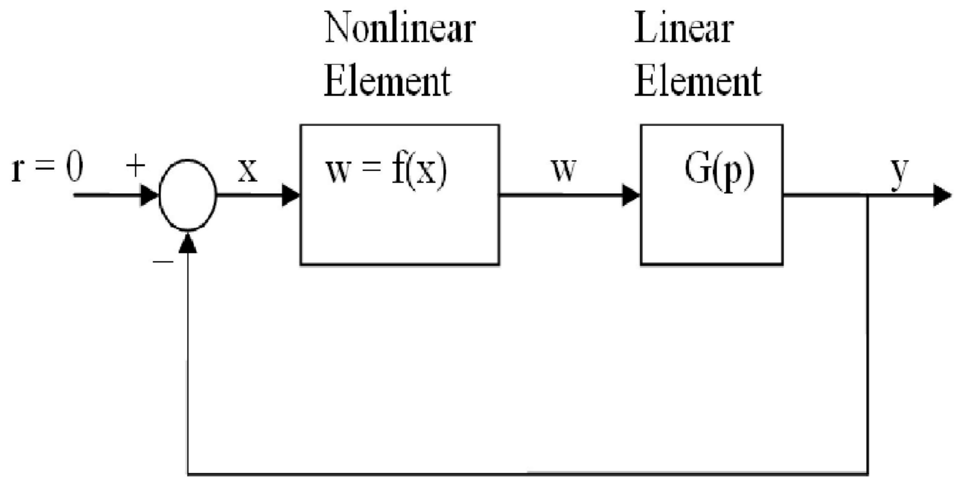


Figure 1:

systems, we refer to systems which contain all linear block except for one nonlinear block in the control loop. Such systems arise when a control system is designed using linear control but its implementation involves hard nonlinearities, such as motor saturation, actuator or sensor dead-zone, Coulomb friction, or hysteresis in the plant.

Example

The plant is linear and the controller is also linear. However, the actuator involves hard nonlinearity. This system can be rearranged by regarding $G_pG_1G_2$ as the linear component G and the actuator nonlinearity as the nonlinear element N.

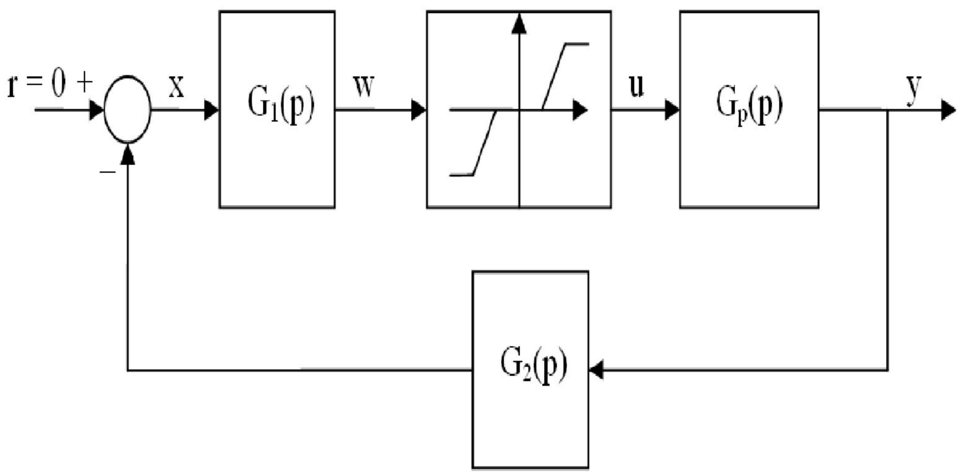


Figure 2:

The applicability to limit cycle analysis is due to the fact that the form of the signals in a limit-cycling system is usually approximately sinusoidal. Consider a sinusoidal input to the nonlinear element, of amplitude A and frequency ω, i.e. $e(t) = A\sin(\omega t)$, as shown in the Figure below

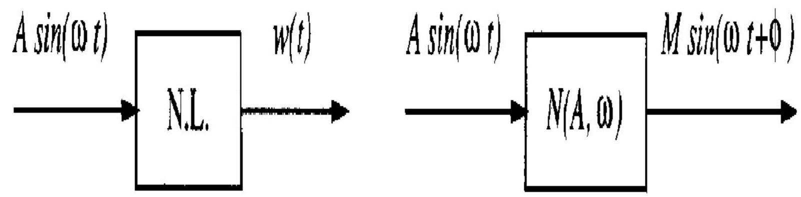


Figure 3:

The output of the nonlinear component $c(t)$ is often a periodic function. Using Fourier series, this periodic function can be expanded as

$$c(t) \approx a_1 \cos(\omega t) + b_1 \sin(\omega t) = M \sin(\omega t + \phi)$$

where

$$a_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} c(t) \cos(\omega t) d(\omega t) \tag{1}$$
$$b_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} c(t) \sin(\omega t) d(\omega t) \tag{2}$$
$$M = \sqrt{a_1^2 + b_1^2}$$
$$\phi = \arctan(\frac{a_1}{b_1})$$

In complex representation, this sinusoid can be written as

$$c \approx M \angle \phi$$

Similar to the concept of frequency response function, which is the frequency domain ratio of the sinusoidal input and the sinusoidal output of a system, the describing function of the nonlinear element is the complex ratio of the fundamental component of the input and output signals:

$$N(A,\omega) = \frac{M\angle\phi}{A\angle 0} = \frac{1}{A}(b_1 + ja_1)$$

The nonlinear element, in the presence of sinusoidal input, can be treated as if it were a linear element with a frequency response function $N(A,\omega)$. The describing function, unlike linear system frequency response transfer function, changes with input amplitude A.

Computing Describing Functions

Analytical Calculation

When the nonlinear characteristics $c = f(e)$ are given by an explicit function and the integration in Equations (1) and (2) can be carried out, then analytical evaluation of the describing function based on Equations (1) and (2) is desirable.

Numerical Integration

For nonlinearities whose input-output relationship $c = f(e)$ is given by graphs or tables, numerical integration can be used to evaluate thc dcscribing functions. Thc idca is, of coursc to approximatc thc integral in Equations (1) and (2) by discrete sums over small intervals.

Experimental Evaluation

When a system nonlinearity can be isolated and excited with sinusoidal inputs of known amplitude and frequency, experimental determination of the describing function can be obtained by using a harmonic analyzer on the output of the nonlinear element. This is quite similar to the experimental determination of frequency response functions for linear elements. The difference here is that not only the frequencies, but also the amplitudes of the input sinusoidal should be varied. The results of of the experiments are a set of curves on the complex plane representing the describing function $N(A, \omega)$, instead of analytical expressions.

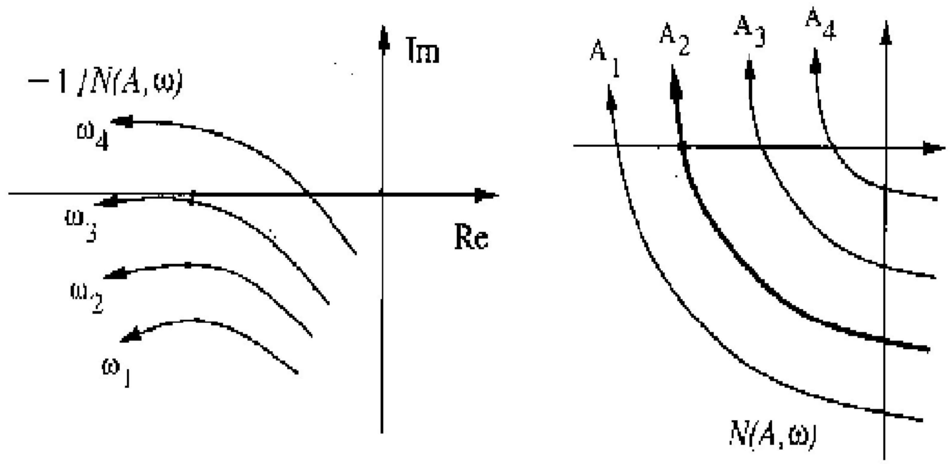


Figure 5:

Describing Functions of Common Nonlinearities

Saturation

$$c(t) = \begin{cases} kA\sin(\omega t) & 0 \le \omega t \le \omega t_1 \\ ka & \omega t_1 < \omega t \le \pi/2 \end{cases}$$

where $\omega t_1 = \sin^{-1}(a/A)$. The odd nature of $c(t)$ implies that $a_1 = 0$ and the symmetry over the four quarters of a period implies that

$$\begin{aligned} b_1 &= \frac{4}{\pi}\int_0^{\pi/2} c(t)\sin(\omega t)d(\omega t) \\ &= \frac{4}{\pi}\int_0^{\omega t_1} kA\sin^2(\omega t)d(\omega t) + \frac{4}{\pi}\int_{\omega t_1}^{\pi/2} ka\sin(\omega t)d(\omega t) \\ &= \frac{2kA}{\pi}[\omega t_1 + \frac{a}{A}\sqrt{1 - \frac{a^2}{A^2}}] \end{aligned}$$

The describing function is $N(A) = \frac{b_1}{A}$.

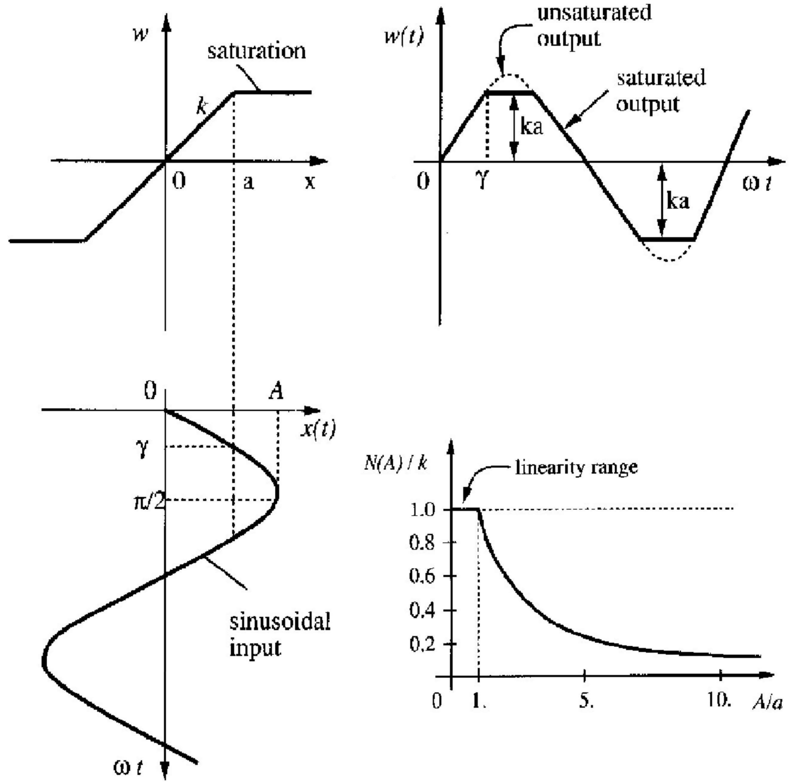


Figure 6:

Relay

$$c(t) = \begin{cases} M & 0 \leq \omega t \leq \pi \\ -M & \pi < \omega t \leq 2\pi \end{cases}$$

$$b_1 = \frac{4}{\pi} \int_0^{\pi/2} M \sin(\omega t) d(\omega t) = \frac{4}{\pi} M$$

Describing function:

$$N(A) = \frac{4M}{\pi A}$$

Existence of Limit Cycle

Similar to Linear system condition for sustained oscillation, nonlinear

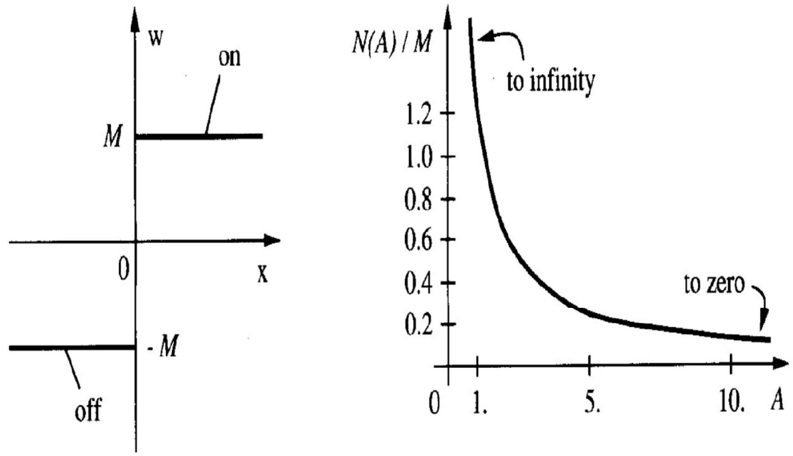


Figure 7:

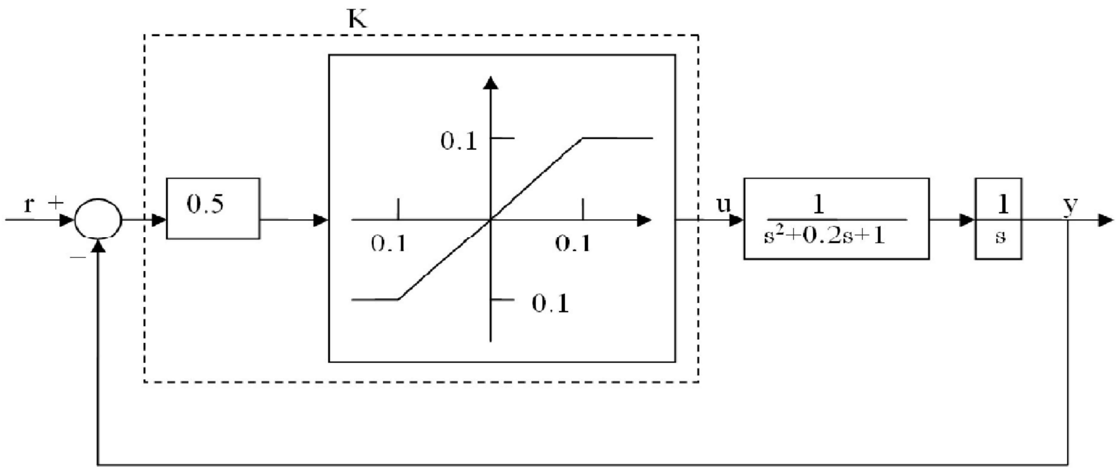
system limit cycling is obtained when

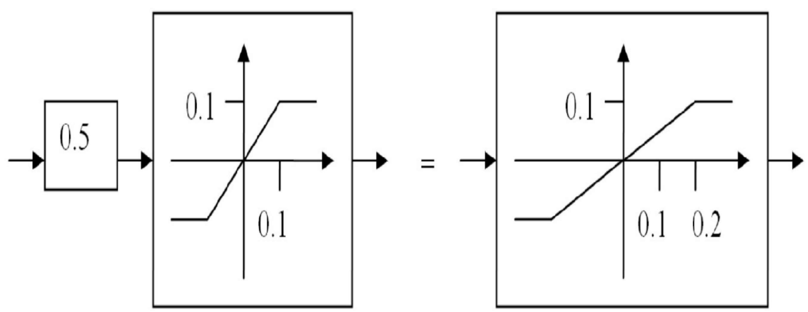
$$G(j\omega)N(A,\omega) = -1$$

which can be written as

$$G(j\omega) = -\frac{1}{N(A,\omega)}$$

Example





$$
\begin{aligned}
N(A) &= \frac{2\times k}{\pi}\left[\sin^{-1}\left(\frac{a}{A}\right)+\frac{a}{A}\sqrt{1-\left(\frac{a}{A}\right)^2}\right] \\
&= \frac{2\times 0.5}{\pi}\left[\sin^{-1}\left(\frac{0.2}{A}\right)+\frac{0.2}{A}\sqrt{1-\left(\frac{0.2}{A}\right)^2}\right]
\end{aligned}
$$

$$
\begin{aligned}
G(j\omega) &= \frac{1}{j\omega(-\omega^2+0.2j\omega+1)} \\
N(A)G(j\omega) &= -1
\end{aligned}
$$

$$
\frac{1}{\pi}\left[\sin^{-1}\left(\frac{0.2}{A}\right)+\frac{0.2}{A}\sqrt{1-\left(\frac{0.2}{A}\right)^2}\right] = j\omega^3+0.2\omega^2-j\omega
$$

Compare Imaginary Part

$$
\omega^3-\omega = 0; \qquad \omega = 0 \text{ or } 1
$$

Compare Real Part

$$0.2 = \frac{1}{\pi}\left[\sin^{-1}\left(\frac{0.2}{A}\right) + \frac{0.2}{A}\sqrt{1-\left(\frac{0.2}{A}\right)^2}\right] \qquad A = 0.626$$

LINEARIZATION

Gain scheduling is an attempt to apply the well developed linear control methodology to the control of nonlinear systems. The idea of gain scheduling is to select a number of operating points which cover the range of the system operation. Then, at each of these points, the designer makes a linear time-invariant approximation to the plant dynamics and designs a linear controller for each linearized plant. Between operating points, the parameters of the compensators are then interpolated.

Consider the cascaded spherical tanks below.

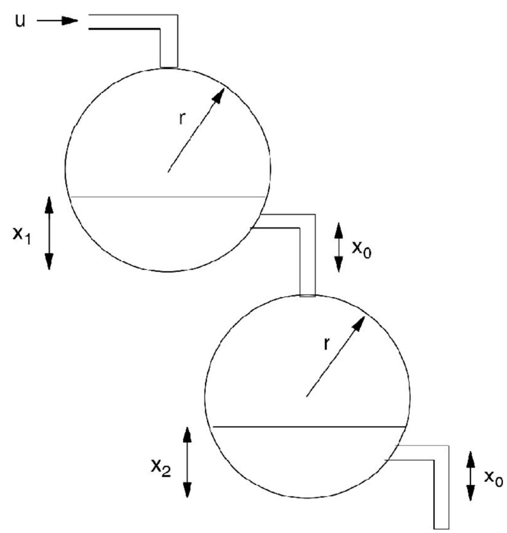


Figure 1: Spherical Tank

The rate of change in volume is given by

$$
\begin{aligned}
v &= \frac{1}{3}\pi x^2(3r - x) \\
\frac{dv}{dt} &= \pi(2rx - x^2)\dot{x}
\end{aligned}
$$

This must be equal to the rate of nett inflow. Therefore

$$
\begin{aligned}
u - K\sqrt{x_1 - x_0} &= \pi(2rx_1 - x_1^2)\dot{x}_1 \\
K\sqrt{x_1 - x_0} - K\sqrt{x_2 - x_0} &= \pi(2rx_2 - x_2^2)\dot{x}_2
\end{aligned}
$$

Rearrange and name the two equations as g and h.

$$
\begin{aligned}
\dot{x}_1 &= \frac{u - K\sqrt{x_1 - x_0}}{\pi(2rx_1 - x_1^2)} = g \\
\dot{x}_2 &= \frac{K\sqrt{x_1 - x_0} - K\sqrt{x_2 - x_0}}{\pi(2rx_2 - x_2^2)} = h
\end{aligned}
$$

From Taylor expansion:

$$\begin{aligned}
\dot{x}_1 &= g|_{\bar{x}_1,\bar{x}_2,\bar{u}} + \left.\frac{dg}{dx_1}\right|_{\bar{x}_1,\bar{x}_2,\bar{u}} \Delta x_1 + \left.\frac{dg}{dx_2}\right|_{\bar{x}_1,\bar{x}_2,\bar{u}} \Delta x_2 + \left.\frac{dg}{du}\right|_{\bar{x}_1,\bar{x}_2,\bar{u}} \Delta u \\
\dot{x}_2 &= h|_{\bar{x}_1,\bar{x}_2,\bar{u}} + \left.\frac{dh}{dx_1}\right|_{\bar{x}_1,\bar{x}_2,\bar{u}} \Delta x_1 + \left.\frac{dh}{dx_2}\right|_{\bar{x}_1,\bar{x}_2,\bar{u}} \Delta x_2 + \left.\frac{dh}{du}\right|_{\bar{x}_1,\bar{x}_2,\bar{u}} \Delta u
\end{aligned}$$

At steady state

$$\begin{aligned}
g|_{\bar{x}_1,\bar{x}_2,\bar{u}} &= h|_{\bar{x}_1,\bar{x}_2,\bar{u}} = 0 \\
g|_{\bar{x}_1,\bar{x}_2,\bar{u}} &= \frac{\bar{u} - K\sqrt{\bar{x}_1 - x_0}}{\pi(2r\bar{x}_1 - \bar{x}_1^2)} = 0 \Rightarrow \bar{u} = K\sqrt{\bar{x}_1 - x_0} \\
h|_{\bar{x}_1,\bar{x}_2,\bar{u}} &= \frac{K(\sqrt{\bar{x}_1 - x_0} - \sqrt{\bar{x}_2 - x_0})}{\pi(2r\bar{x}_2 - \bar{x}_2^2)} = 0 \Rightarrow \bar{x}_1 = \bar{x}_2 = (\frac{\bar{u}}{K})^2 + x_0
\end{aligned}$$

The linearized model in state space form

$$\begin{bmatrix} \dot{\Delta} x_1 \\ \dot{\Delta} x_2 \end{bmatrix} = \begin{bmatrix} \frac{dg}{dx_1} & \frac{dg}{dx_2} \\ \frac{dh}{dx_1} & \frac{dh}{dx_2} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} + \begin{bmatrix} \frac{dg}{du} \\ \frac{dh}{du} \end{bmatrix} \Delta u = A\Delta x + B\Delta u$$

$$\Delta y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = C\Delta x$$

where

$$\begin{aligned}
\dot{\Delta} x_1 &= \dot{x}_1 - g|_{\bar{x}_1, \bar{x}_2, \bar{u}} \\
\dot{\Delta} x_2 &= \dot{x}_2 - h|_{\bar{x}_1, \bar{x}_2, \bar{u}} \\
\frac{dg}{du} &= \frac{1}{\pi(2r\bar{x}_1 - \bar{x}_1^2)} \\
\frac{dg}{dx_1} &= -\frac{dh}{dx_1}
\end{aligned}$$

$$
\begin{aligned}
\frac{dg}{dx_2} &= 0 \\
\frac{dh}{du} &= 0 \\
\frac{dh}{dx_1} &= \frac{K(\bar{x}_1 - \bar{x}_0)^{-1/2}}{2\pi(2r\bar{x}_2 - \bar{x}_2^2)} \\
\frac{dh}{dx_2} &= -\frac{dh}{dx_1}
\end{aligned}
$$

The linearized transfer function model

$$
\frac{\Delta Y(s)}{\Delta U(s)} = C(sI - A)^{-1}B = \frac{\frac{2}{K}\sqrt{\bar{x}_2 - x_0}}{(\frac{2}{K}\pi\bar{x}_2(2r - \bar{x}_2)\sqrt{\bar{x}_2 - x_0}s + 1)^2}
$$

The DC gain $= \frac{2}{K}\sqrt{\bar{x}_2 - x_0}$
The Time constant $= \frac{2}{K}\pi\bar{x}_2(2r - \bar{x}_2)\sqrt{\bar{x}_1 - x_0}$

The PID controller can be designed using pole-placement as follows. Consider the second-order model process

$$G_p(s) = \frac{K_p}{(1+sT_1)(1+sT_2)}$$

The PID controller can be written as

$$G_c(s) = \frac{K_c(1+sT_i+s^2T_iT_d)}{sT_i}$$

The characteristic equation of the closed-loop system becomes

$$s^3 + s^2\left(\frac{1}{T_1}+\frac{1}{T_2}+\frac{K_pK_cT_d}{T_1T_2}\right) + s\left(\frac{1}{T_1T_2}+\frac{K_pK_c}{T_1T_2}\right) + \frac{K_pK_c}{T_iT_1T_2} = 0$$

A suitable closed-loop characteristic equation is a third-order system is

$$(s+\alpha\omega)(s^2+2\zeta\omega s+\omega^2) = 0$$

Comparing coefficients in the last 2 equations gives

$$
\begin{aligned}
K_c &= \frac{T_1T_2\omega^2(1+2\zeta\alpha)-1}{K_p} \\
T_i &= \frac{T_1T_2\omega^2(1+2\zeta\alpha)-1}{T_1T_2\alpha\omega^3} \\
T_d &= \frac{T_1T_2\omega(\alpha+2\zeta)-T_1-T_2}{\omega^2T_1T_2(1+2\zeta\alpha)-1}
\end{aligned}
$$

where α, ζ and ω are user-specified.

SLIDING CONTROL

This is a special version of on-off control. The key idea is to apply strong control action when the system deviates from the desired behavior.

Assume that the system we want to control is described by the nonlinear equation

$$\frac{d^n y}{dt^n} = f_1(y, \frac{dy}{dt}, \ldots, \frac{d^{n-1}y}{dt^{n-1}}) + g_1(y, \frac{dy}{dt}, \ldots, \frac{d^{n-1}y}{dt^{n-1}})u \tag{1}$$

Select the states

$$x = [\, x_1 \; x_2 \; \ldots \; x_{n-1} \; x_n \,]^T = [\, \tfrac{d^{n-1}y}{dt^{n-1}} \; \tfrac{d^{n-2}y}{dt^{n-2}} \; \ldots \; \tfrac{dy}{dt} \; y \,]^T \tag{2}$$

and rewrite the system as

$$
\begin{aligned}
\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdots \\ \dot{x}_n \end{bmatrix} &= \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} + \begin{bmatrix} f_1(x) + g_1(x)u \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} f_1(x) \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} + \begin{bmatrix} g_1(x) \\ 0 \\ \vdots \\ 0 \end{bmatrix} u \\
&= f(x) + g(x)u \qquad (3) \\
y &= [\, 0 \;\; 0 \;\; \dots \;\; 0 \;\; 1 \,] x \qquad (4)
\end{aligned}
$$

where $f(x)$ and $g(x)$ are vectors.

Let the switching surface be

$$\sigma(x) = p_1 x_1 + p_2 x_2 + \cdots + p_n x_n = p^T x = 0$$

Using the definition of the state vector

$$\sigma(x) = p_1 \frac{d^{n-1}y}{dt^{n-1}} + p_2 \frac{d^{n-2}y}{dt^{n-2}} + \cdots + p_n y = 0 \tag{5}$$

The dynamic behavior on the sliding surface can be specified by a proper choice of the number p_i. It will be stable if the polynomial

$$P(s) = p_1 s^{n-1} + p_2 s^{n-2} + \cdots + p_n \tag{6}$$

has all its roots in the left-half plane.

To determine a control law that keeps the system on $\sigma(x) = 0$, we introduce the Lyapunov function

$$V(x) = \sigma^2(x)/2 \tag{7}$$

The time derivative of V is given by

$$\frac{dV}{dt} = \sigma(x)\dot{\sigma}(x) = \sigma p^T \dot{x} = \sigma(p^T f(x) + p^T g(x)u(t)) \tag{8}$$

Choose the control law

$$u(t) = -\frac{p^T f}{p^T g} - \frac{\mu}{p^T g}\text{sign}(\sigma(x)) \tag{9}$$

so that

$$\frac{dV}{dt} = -\mu\sigma(x)\text{sign}(\sigma(x)) \tag{10}$$

is always negative for $\sigma(x) \neq 0$.

Steady State Response

This must mean that at steady state ($t \to \infty$): (i) from Equations (7) and (10), $\sigma(x) = 0$, (ii) from Equation (5) y = 0.

Transient Response

Assume that the system has initial values $\sigma(x) = \sigma_0 > 0$, and let t_σ be the time when the switching surface is reached. From Equations (8) and (9) we find that

$$\dot{\sigma}(x) = -\mu$$

Integrating this equation

$$\int_{\sigma_0}^{0} d\sigma(x) = -\int_0^{t_\sigma} \mu dt$$

gives

$$0 - \sigma_0 = -\mu(t_\sigma - 0)$$

which gives $t_\rho = \sigma_0/\mu$. Using the same arguments for $\sigma_0 < 0$ shows that $t_\sigma = |\sigma_0|/\mu$. The subspace $\sigma(x) = 0$ is asymptotically stable, and the state will stay on the switching surface once it is reached. The motion along the surface is determined by Equation (5).

Uncertainties in f and g can be handled if μ is sufficiently large. Assume that the design of the control law is based on the approximate values $\hat{f}$ and $\hat{g}$ instead of the true ones. Then

$$\frac{dV}{dt} = \sigma \left(\frac{p^T(f\hat{g}^T - \hat{f}g^T)p}{p^T\hat{g}} - \mu \frac{p^T g}{p^T \hat{g}} \operatorname{sign}(\sigma) \right)$$

The right-hand side is negative if μ is sufficiently large, provided that $p^T\hat{g}$ and $p^T g$ have the same sign. The system will thus be insensitive to uncertainties in the process model.

<u>Smooth Control Laws</u>

The control law (9) has the drawback that the relay chatters. One way to avoid this is to make the relay characteristics smoother. The sign function in Equation (9) is now replaced by the saturation function

$$\mathrm{sat}(\sigma, \varepsilon) = \begin{cases} 1 & \sigma > \varepsilon \\ \sigma/\varepsilon & -\varepsilon \leq \sigma \leq \varepsilon \\ -1 & \sigma < -\varepsilon \end{cases}$$

The control law is then

$$u(t) = \frac{p^T f}{p^T g} - \frac{\mu}{p^T g}\mathrm{sat}(\sigma(x), \varepsilon) \tag{11}$$

<u>Example</u>

Consider the unstable system

$$\frac{dx}{dt} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u = Ax + Bu$$

$$y = \begin{bmatrix} 0 & 1 \end{bmatrix} x$$

which has the transfer function

$$G(s) = \frac{1}{s(s-1)}$$

Choose

$$\sigma(x) = p_1 x_1 + p_2 x_2 = x_1 + x_2 = \frac{dy}{dt} + y = 0$$

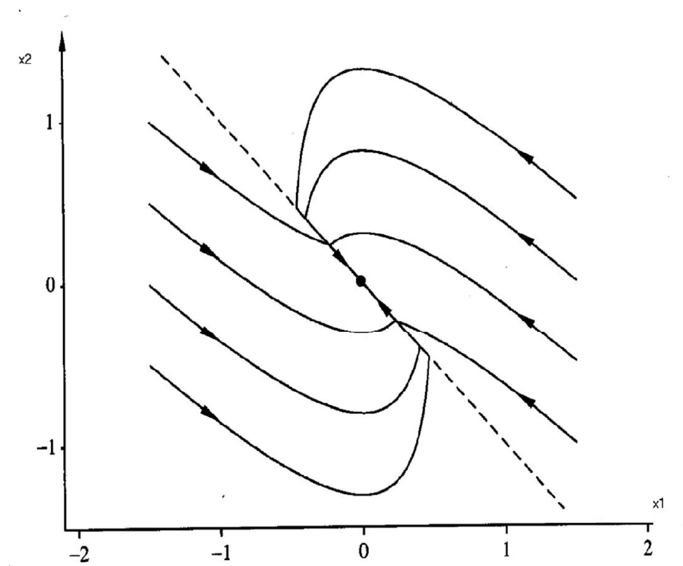
The controller from Equation (9) is now

$$\begin{aligned} u(t) &= -\frac{p^T A x}{p^T B} - \mu\,\mathrm{sign}(\sigma(x)) \\ &= -[\,2\ \ 0\,]x(t) - \mu\,\mathrm{sign}(\sigma(x)) \end{aligned}$$

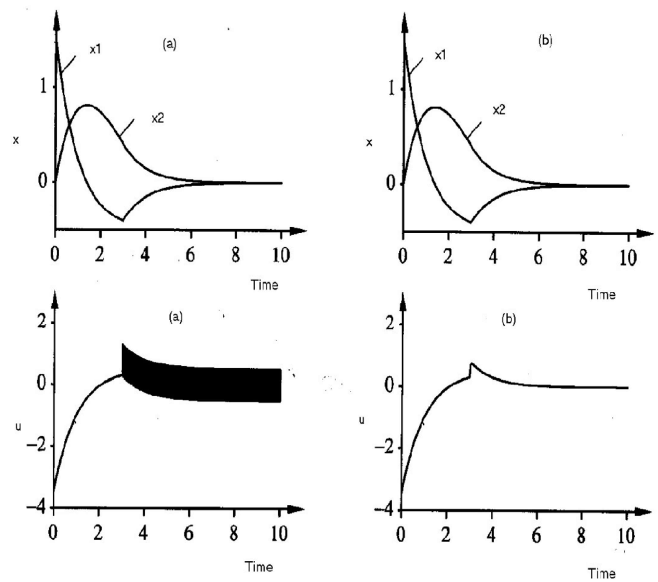
The smooth control law is

$$u(t) = -[\,2\ \ 0\,]x(t) - \mu\,\mathrm{sat}(\sigma(x))$$

The phase plane when $\mu = 0.5$ is shown below. The input and output for both control laws are also shown below.



Phase portrait of the system in the Example. The dashed line shows $\sigma(x) = 0$.



The states and the output as a function of time in the Example. The initial conditions are $x_1(0) = 1.5$ and $x_2(0) = 0$. The controllers are (a) Eq. (9) with $\sigma = 0.5$; (b) Eq. (11) with $\sigma = 0.5$ and $\varepsilon = 0.01$.